Datacomputer and SIP Operations: 1978 Final Technical Report

LEVEL II

D D C
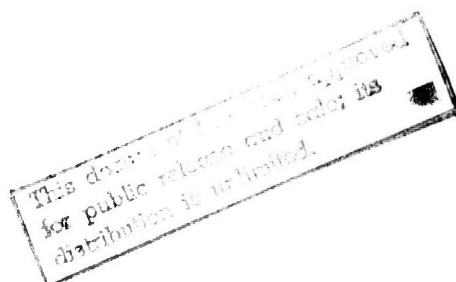MAY 15 1979

Technical Report
CCA-79-11
February 26, 1979

Donald E. Eastlake, III
Matthew Maltzman
Joanne Z. Sattley

Datacomputer and SIP
Operations

1978 Final Technical
Report for period ending 1978

Technical Report
CCA-79-11

Donald E. Eastlake, III
Matthew Maltzman
Joanne Z. Sattley

30 March, 1979

387285

Table of Contents

List of Figures

List of Tables

1.   Summary

During 1978, Computer Corporation of America offered very
large on-line data storage and retrieval services on the
Datacomputer in support of the seismic community data
activities and for general use. The volume of data
stored, the number of users, and the diversity of user
projects was substantial at the start of the year and
continued to grow throughout the reporting period.

The Datacomputer is a system designed to allow convenient
and timely access to large on-line databases for multiple
remote users communicating over a network. The
Datacomputer is the only operational general purpose
database system capable of handling data sets in excess of
a trillion bits.

Copious and inexpensive storage is a unique feature of the
CCA Datacomputer, made possible by the incorporation of an
Ampex Tera-Bit Memory System (TBM). The TBM at CCA was
the first public installation of this video-tape
technology based system. The CCA installation is
configured to hold up to 175 billion bits on-line with
four TBM tape drives. Additional data (almost entirely
seismic) is stored in off-line TBM tapes.

Appendix A to this report gives an overview of the
Datacomputer itself; Appendix B is an historical account
of the development of the Datacomputer.

In addition to operating the Datacomputer, CCA also
provides help and consultation to both its seismic and
general user community. This assistance frequently takes
the form of technical advice which is aimed at achieving
the maximum benefit from the unique capabilities offered
by the Datacomputer. As another form of assistance, CCA
has implemented and continues to maintain several programs
and subroutines which run on remote Arpanet user hosts;
these provide users across the Arpanet with convenient
interfaces to the Datacomputer. DFTP and DCSTAT are
prominent examples of such popular utility programs.

The development of a novel Message Archiving and Retrieval
Service (MARS) has provided yet another use for the large
storage capacity and lookup capabilities of the
Datacomputer. MARS users do not communicate directly with
the Datacomputer, but rather, use ordinary network mail
channels to transmit messages for filing and retrieving;
program demons operate on CCA-Tenex to perform and manage
the Datacomputer interactions.

Particular effort was directed in 1978 toward assisting in
the design for the Signal Waveform File project, and

toward implementation by CCA of a part of this new seismic application for the Datacomputer.

The seismic application is the largest user of the Datacomputer. This is true not only in terms of the amount of data stored, but also in terms of its complexity, and the bandwidth of transfer. Some of the data involved is sent in real time to CCA, but not directly to the Datacomputer. The real-time data stream is fielded by a small reliable dedicated processor, called the Seismic Input Processor, or SIP. This sturdy interface was designed and implemented by CCA; the SIP reformats received data for efficiency, and periodically forwards it in quantity to the Datacomputer.

Maintenance of the Datacomputer itself is continuing on a regular basis in conjunction with a modest amount of work on desirable software improvements. Some of the developmental work is being funded under a separate contract (N00039-77-C-0074) -- but its fruits are made available to all Datacomputer users. 1978 also saw the issuance of an all-new complete Datacomputer user manual and other user-oriented documentation.

The CCA Datacomputer runs under a TENEX operating system which has been extensively modified to accommodate it. During 1978, some additional modifications to the system

and to one of its utility programs were made in order to
facilitate the continuing increase in the size of the
Datacomputer directory and to enable faster response to
occasional equipment malfunctions. Other general support
software upgrades were also made.   Some problems were
encountered with the CCA-TENEX Arpanet interface due to
work on other hosts connected through our on-site PLURIBUS
IMP and the fact that this IMP, though a multi-processor,
has only one bus.

Overall, the Datacomputer and SIP continue to provide
unique very large database management services to seismic
and other users over the Arpanet.

2. Support of the User Community

One of CCA's most important tasks under this contract is
to provide assistance and coordination for the seismic
community and for other users of the Datacomputer. Such
assistance and coordination are not only beneficial but
necessary to users so that they can get the most out of
the Datacomputer as a network utility resource.

Below we give overviews of the seismic and general user
communities, details on the Datacomputer operations and
usage trends in 1978, and descriptions of significant
developments in the CCA-provided Datacomputer user
interface DCLINK subroutine package, DFTP program, and
MARS service.

2.1 The Seismic User Community

The seismic users of the Datacomputer (see Table 2.1) are

primarily concerned with research in the seismic nuclear

Seismic Datacomputer Users   Table 2.1

Albuquerque Seismological Laboratory

Lincoln Laboratory
        Applied Seismology Group

National Earthquake Information Service

Seismic Data Analysis Center
        Teledyne Geotech
        Texas Instruments

Vela Seismological Center
----

monitoring area. Their principal source of data is the

worldwide network of seismic instruments, transmission

lines, and data processors known as the Vela Network. Some

components of the Velanet are on the Arpanet and use it as

a transmission system, while other parts use leased lines

for real time data or the shipment of magnetic tape for

non-real time data. Figure 2.1 shows the Vela Network

configuration superimposed on a map on the Arpanet. Figure

2.2 (provided by the Seismic Data Analysis Center (SDAC))

Seismic Data Network
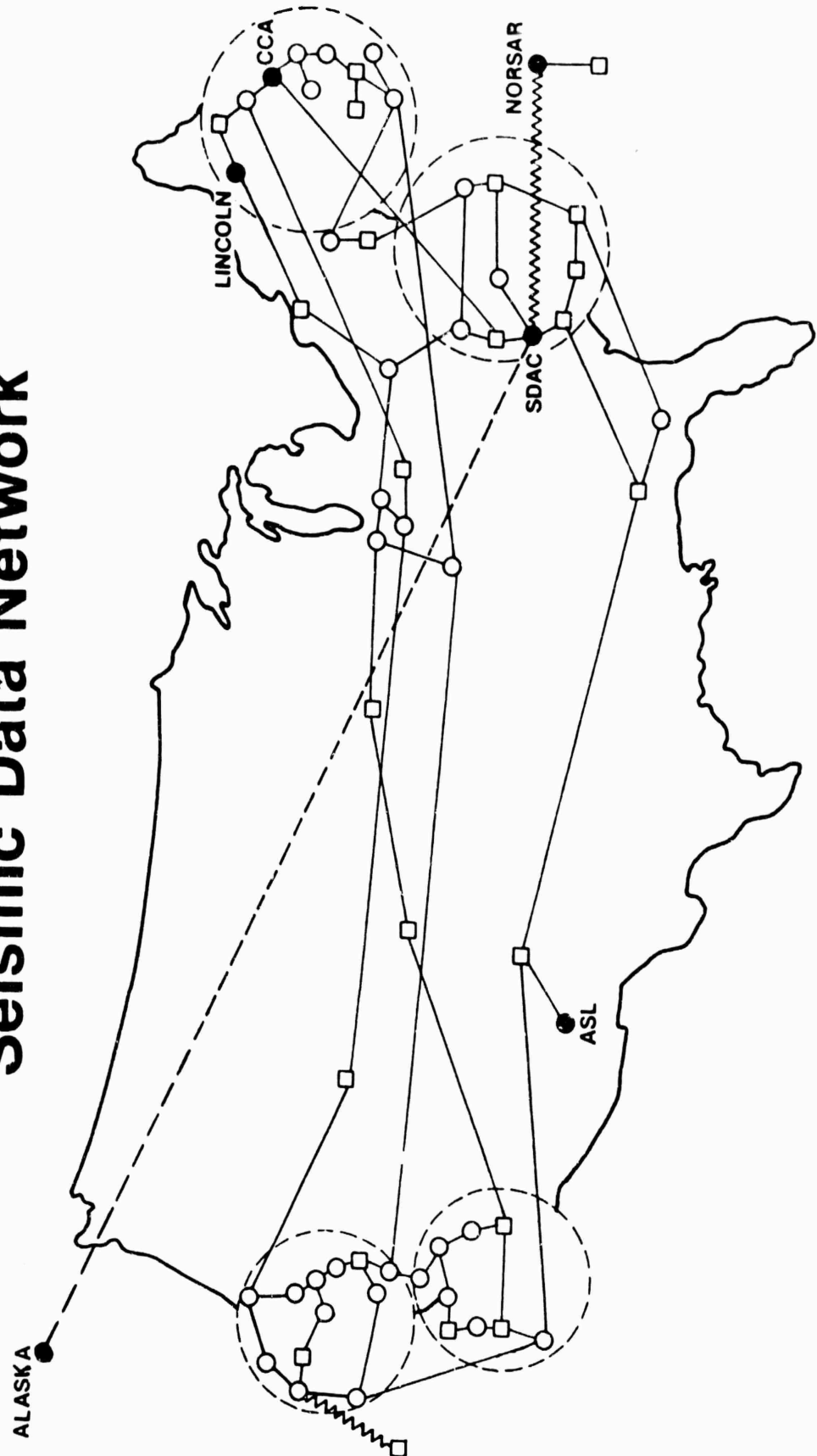
shows the geographic dist. bution of Velanet instruments
and other key components.

The Datacomputer at CCA is the primary storage and
retrieval resource in the Velanet. The seismic data
activity requires storage of very large amounts of online
data. This data includes the following·

1. seismic readings from arrays of instruments, some
    of which are sent in real time;

2. seismic readings from individual instruments which
    are forwarded from the Albuquerque Seismological
    Laboratory by the US Geological Survey;

3. limited status and calibration information on
    instruments and the Velanet;

4. derived seismic event summary information; and

5. extracted signal waveforms corresponding to events.

Seismic instrument readings can be further subdivided into
long period (one sample per second) and short period (ten
or twenty samples per second) data.

Since the Datacomputer is not designed to receive data in
real time, a special dedicated miniprocessor, the SIP, is
used to buffer the real time seismic instrument array
data. (Section 5 below discusses the SIP.) The status
and event summary data are relatively compact and are of

Geographic Representation of VELA Network   Figure 2.2

LEGEND

★ -SDAC

◣ -MASS STORE

● -REAL-TIME STATIONS

□ -SRO/HGLP STATIONS
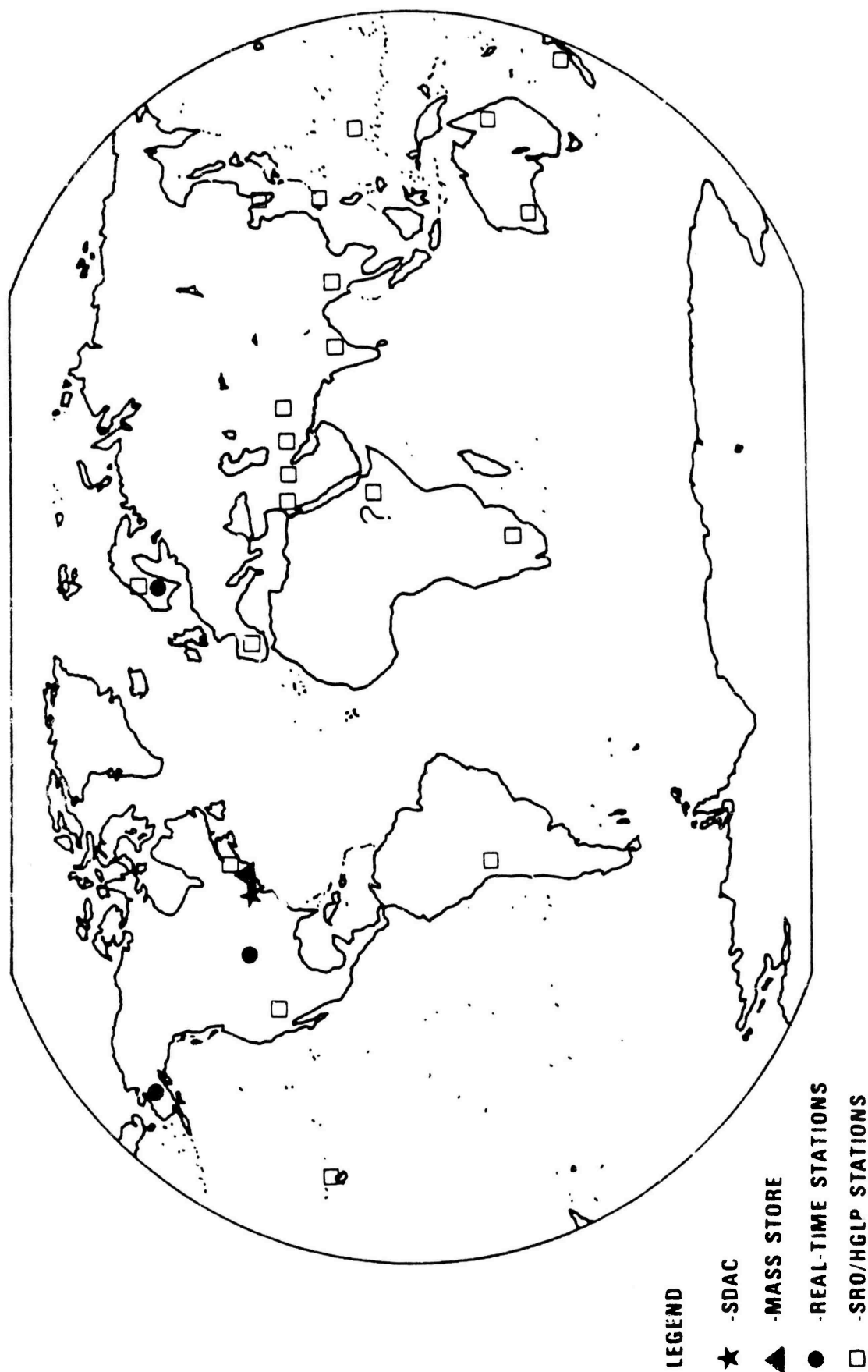
sufficient importance that they are kept on-line indefinitely. Seismic data from single instruments and arrays of instruments is sufficiently voluminous that it fills many reels of tape, most of which are kept off line, in our mass storage system. (Section 6 below discusses the TBM mass storage system in detail.) Seismic Signal Waveform data is the most recent type to be stored in the Datacomputer and is covered in Section 4 below. Figure 2.3 shows the overall accumulation of seismic data in the Datacomputer through the end of 1978.

2.2 The General User Community

The general Datacomputer user community includes a large number and variety of non-seismic users. Some of the organizations from which these users come are listed in Table 2.2.

The amount of information that had been stored in the Datacomputer by non-seismic users, through the end of 1978, was such that it was possible to retain all such information on-line. Figure 2.4 charts the growth in this non-seismic data. Given the amount of seismic data which it is also desirable to retain on line, we expect that in

Megabits

Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

1977   1978

Advanced Command and Control Architectural Testbed (ACCAT)

Air Force Armament Development and Test Center

Air Force Avionics Laboratory

Anniston Army Depot

ARPA

Bolt, Beranek, and Newman, Inc.
        Arpanet Network Control Center

Brookhaven National Laboratory

Lawrence Berkeley Laboratory

Carnegie-Mellon University

Computer Corporation of America
        Message Archiving and Retrieval Service (MARS)

Department of Energy
        Argonne National Laboratory

Digital Equipment Corporation
        Federal Systems Group

Gagliardi Systems Group, Inc.

Harvard University
        ECL Group
        Graphics Group

Intermetrics

Massachusetts Institute of Technology
        Artificial Intelligence Laboratory
        Center for Theoretical Physics
        Laboratory for Computer Science
                Programming Technology Division
                Database Systems Division
        Laboratory for Nuclear Science
        Matnematics Department
        Plasma Fusion Center

Datacomputer and SIP FTR   Page -13-
Support of the User Community   Section 2

Some Non-Seismic Datacomputer Users   Table 2.2
(continued from previous page)

National Bureau of Standards

National Software Works

Rome Air Development Center

Rutgers University

SRI International

Stanford University
        Artificial Intelligence Laboratory
        Stanford University Medical Center

Thames Polytechnic (London, England)

U.S. Army Development and Readiness Command

University of California at Los Angeles

University of Southern California
        Information Sciences Institute

University of Texas at Austin
        Linguistics Research Center

University College London
        Facsimile Transmission Research

White Sand Missile Range

1979 some low-priority non-seismic data will have to be
moved off-line.

Among the significant events for all Datacomputer users in
1978 was the issuance of the completely new Datacomputer
Version 5 User Manual and Datacomputer Technical Bulletin
Number 8 on the Datacomputer status server. Both these
documents are described in Section 3 below.

Datacomputer and SIP FTR
Support of the User Community

Page -14-
Section 2

Non-Seismic Data Accumulated

Figure 2.4

MEGABITS
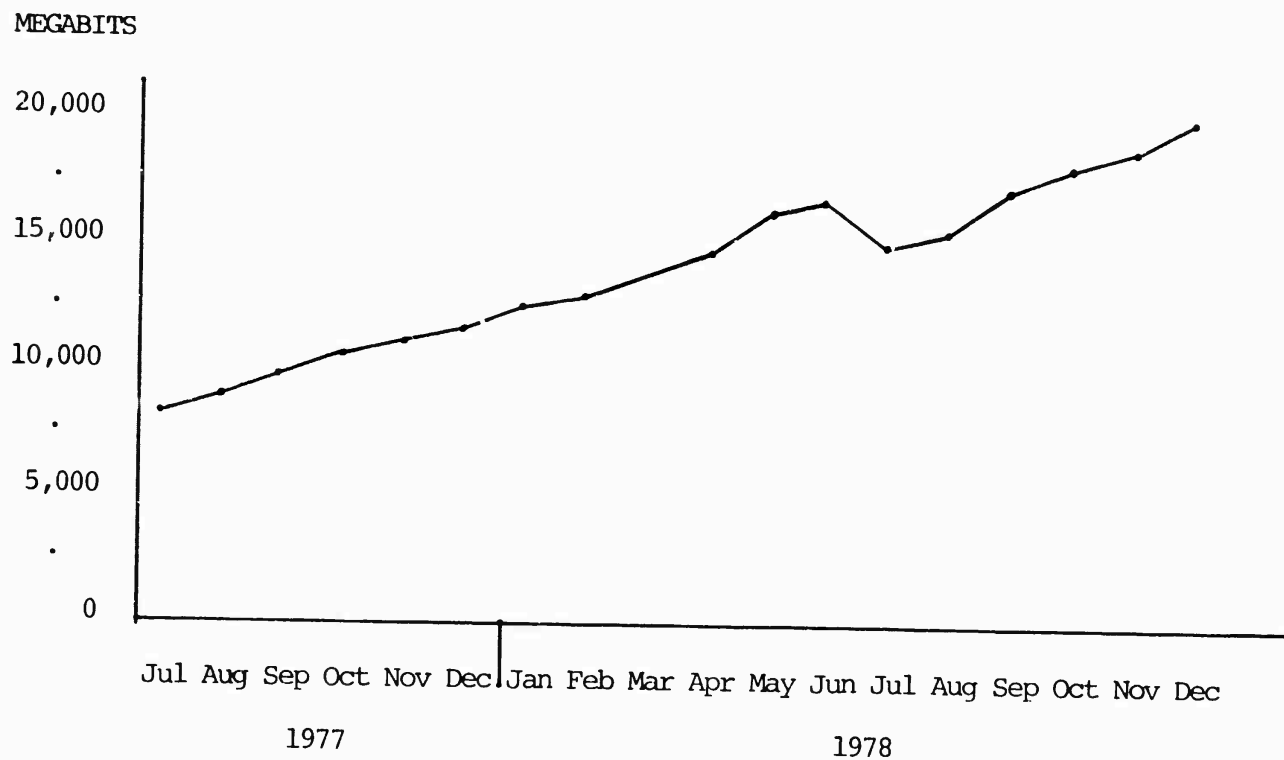
Figure 2.4

The small interval in Figure 2.4 where the amount of non-seismic data declines was due to CCA assistance to the Bolt, Beranek and Newman IMP LOGGER application. This Datacomputer application stores statistics and event information related to the Arpanet communications backbone in the Datacomputer. Initially it used daily files which were relatively small and inefficient and effectively limited the amount of information that could be kept on-line. The application was converted by CCA to use more efficient monthly files and the decline in Figure 2.4 represents the deletion of the old daily files after the new version of the application was operational.

Among the new Datacomputer users that were contacted for the first time in 1978 were the following: linguist Bob Amsler of the University of Texas, the ACCAT GUARD project at Logicon, the ARPA S1 project, T. Bowerman of the Anniston Army Depot, Tony Michel of BBN, COOPRIDER@ISIB, and Frank Wancho of the White Sands Missile Range.

2.3 Operations and Usage Trends

Figure 2.5 shows the availability of the Datacomputer during prime time (9AM to 8PM business days) (the dotted lines in February show the effect of including the time of the Great Blizzard of 1978 in these figures). Table 2.3 gives a breakdown of the causes of unavailability into several broad classes. It will be seen that almost a factor of two improvement in reducing unavailability has occurred from the later half of 1977 to the later half of 1978.

To enable users with problems to contact a member of CCA's staff, we maintain a Datacomputer trouble telephone number and an Arpanet address (HELP@CCA) to which trouble messages can be sent. The trouble telephone number rings at an answering service that can take trouble messages and beep someone who will respond.

To automatically alert CCA personnel more rapidly in the case of certain types of hardware problems, a background process in the CCA-TENEX system called ALL was modified as described in Section 7.1 below.

Datacomputer and SIP FTR
Support of the User Community

Page -18-
Section 2

Datacomputer Availability Statistics

Table 2.3
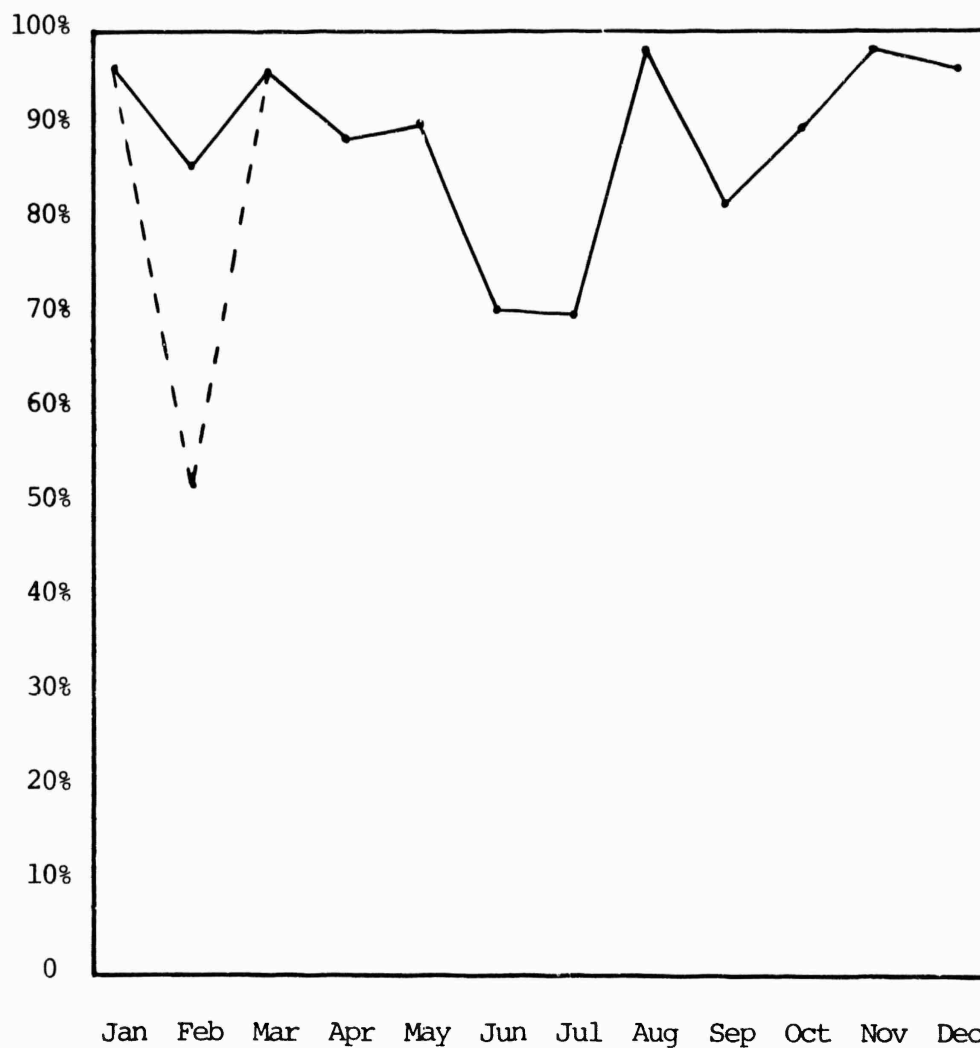
<u>Datacomputer Prime Time Unavailability</u>

Jul - Dec 1977	-	262.8 hours
Jan - Jun 1978	-	193.0 hours
Jul - Dec 1978	-	136.2 hours

Breakdown by Cause of Unavailability

<u>Cause</u>	<u>JulDec77</u>	<u>JanJun78</u>	<u>JulDec78</u>	<u>Overall</u>
TBM	65.8	42.5	31.7	50.4
Tenex/DC	27.6	44.3	30.8	33.8
Network	4.6	7.5	28.3	11.0
Operator	1.1	4.1	1.1	2.1
Environment	0.4	1.0	5.7	1.8
Staging Disk	0.5	0.5	2.4	0.9
	100.0	100.0	100.0	100.0

Usage during 1978 is charted in Figure 2.4. The most
remarkable usage datum over the course of the year was the
tremendous growth of the CCA Message Archiving and
Retrieval Service (MARS) which is described in Section
2.4.2 below.

2.4 User Interfaces

The development of interfaces to make the Datacomputer's
data storage services more easily accessible is a
significant part of the work done by CCA. The following
sections cover three interfaces on which work was done in
1978. They range from a new subroutine package to a file
storage and retrieval program to a complete message
archiving and retrieval service.

2.4.1 DCLINK

DCLINK is a new subroutine package that provides a
convenient Datacomputer interface to programs on TENEX and
TOPS-20 systems. Written to remedy certain design-level
shortcomings of its predecessor, DCSUBR, DCLINK generally
embodies improvements in functionality, maintainability,
and ease of use. Data was first transferred successfully
with DCLINK on 29 November 1978. This section will
discuss the motivation for DCLINK's development and then
mention some of its significant features.

The Datacomputer and a remote user communicate over at
least two network connections: one carrying mainly
Datalanguage from user to Datacomputer, another carrying
mainly messages from Datacomputer to user. Separate
connections may be established to handle data transfers.
The Datacomputer occasionally halts a transfer (over
whatever connection) while it generates a message
(describing, e.g., an error or the progress of staging)
and sends this to the user over the message connection.
Since a net connection is like a pipeline of finite
capacity, it can fill up if the sender puts enough in and
the receiver never removes anything. An attempt to write
into a "full" connection will hang as the sending system's
Network Control Program waits for the receiving system to
make some room in the "pipe". One shortcoming of DCSUBR
is that it was designed as a single-process program. It
can therefore perform I/O on only one connection at
time. It may hang when trying to transfer data because
the corresponding Datacomputer job is hung trying to write
into the message connection. This can occur because the
message connection can fill up while DCSUBR is busy
transferring data.

Other factors motivating the replacement of DCSUBR are its
buffering method, which allows the loss of Datacomputer
messages, an obscure programming interface, and the
desirability of new features to support SDD-1.

DCLINK is comprised of multiple independent processes, one
of which is always free to read the message connection,
and thus DCSUBR-like network deadlocks are impossible.
DCLINK's message handling scheme uses temporary disk files
to preclude loss of messages due to core buffer overflow.

Emphasis was placed on making it easy to build application
programs around DCLINK. The definition of individual
routines and their returns represents an attempt to
minimize the differences between DCLINK function semantics
and Datalanguage semantics. For example, DCSUBR's
"connect a Datacomputer port to a local socket" function
is replaced by a broader connection-definition function
that allows a port to be linked not only with a socket but
also with a local data file. Then when the application
sends the Datalanguage to execute a data transfer through
that port, DCLINK automatically mediates the entire
transfer without requiring any other action on the part of
the application program.

A major feature of DCLINK is its acceptance of keyword
parameters. Instead of passing a fixed set of parameters
in registers, application code passes DCLINK the address
of a block containing (keyword,value) pairs and
individual keywords. The block need contain only what is
relevant to a particular call, and the entries need not be

in any particular order. DCLINK validates the contents of
the   rameter block on each call.   The result is
tremendous flexibility in the design of both application
programs and backward-compatible DCLINK modifications,
which is well worth the slight execution overhead in
DCLINK's network-bound operating environment.

Another DCLINK feature will permit exploitation of the
parallelism available in a computer network. Even if they
are not implemented in a multi-process fashion,
application programs will be able to execute their own
code while the Datacomputer is executing their
Datalanguage, and they can catch up with or wait for the
Datacomputer at their convenience.

Some other features representing improvements over DCSUBR
are:

  - support for multiple script files, allowing
    convenient scripting to both a terminal and a more
    permanent file
  - selective suppression of the scripting of any data
    transferred over the Datalanguage or message
    connections

- elimination of redundant buffering during data
  transfers
- automatic classification of Datacomputer messages
  into useful categories, with provision for
  category-based retrieval

To date DCLINK has been incorporated into a version of the
existing program RDC and may be retrofitted elsewhere. It
is expected to be used in SDD-1 and other new Datacomputer
applications.

2.4.2 DFTP: Datacomputer File Transfer Program

DFTP provides a simple interface to the Datacomputer for
the storage and retrieval of whole local files. In terms
of number of individuals making use of a facility, it is
the most popular Datacomputer application.

Several enhancements were made to DFTP in 1978. In
particular, to adapt it better to use from MULTICS systems
and from other systems when trying to reference files
originally stored from MULTICS, the following features
were added: a mode in which upper and lower case
characters are distinguished in filenames; a reasonable,
though limited, way to handle file names with more than

two parts; and several special commands for converting
text to and from the 9-bit-byte MULTICS format.

2.4.3 MARS, a Message Archiving and Retrieval Service

The Datacomputer-based Message Archiving and Retrieval
System, MARS, has been in service since early in 1978. The
System was developed in 1977 under ARPA Contract No.
N00014-76-C-0991, as an application package which would
utilize the Datacomputer [Dorin & Sattley] for long-term
storage of thousands of Arpanet messages and for
convenient retrievals of subsets of these messages. Given
the quickened pace of communications supported on the
Arpanet and the increased volume of electronic message
traffic, the archiving of mail on the Datacomputer
provides a safeguard against the loss of information, and
can alleviate local-site storage requirements for those
who wish to keep a record of their correspondence.

An important element in the concept of paperless
communications is the elimination of file-drawers full of
correspondence and memos. Using the large storage capacity
of the Datacomputer and reliable MARS Service, it is
possible to keep many years' worth of mail and meeting
notes on-line, and available for immediate retrieval. The

extensive indexing performed on the message when it is filed allows it to be retrieved in response to many different requests, as if, in a paper-filing system, many duplicate copies were filed in different folders.

A technical memorandum which describes MARS' usage of the Datacomputer has been distributed as Network Working Group RFC #744, NIC #42827. In April, 1978, an informal report on the first few months operations was distributed to MARS users. [Sattley c]

Since the inception of MARS service, the amount of mail received for archiving has grown to an average of 1000 messages per week; this represents a net monthly saving of 1600-2000 pages of local disk storage for the users of MARS. Netwide users include several teleconferencing groups; in particular, all of the MsgGroup, Header-People, SUPDUP-Group, and EMACS-History messages are available to all Arpanet correspondents. We are also storing Internet-Working-Group and Transmission-Control-Protocol meeting notes as well as summaries of net host traffic.

Maintaining and operating MARS as a service to the Arpanet community is expected to continue, with new features and operational improvements being incorporated as time and resources permit. Figure 2.6 below reflects the growth of the message database during 1978.

thousands

Dec '77   Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
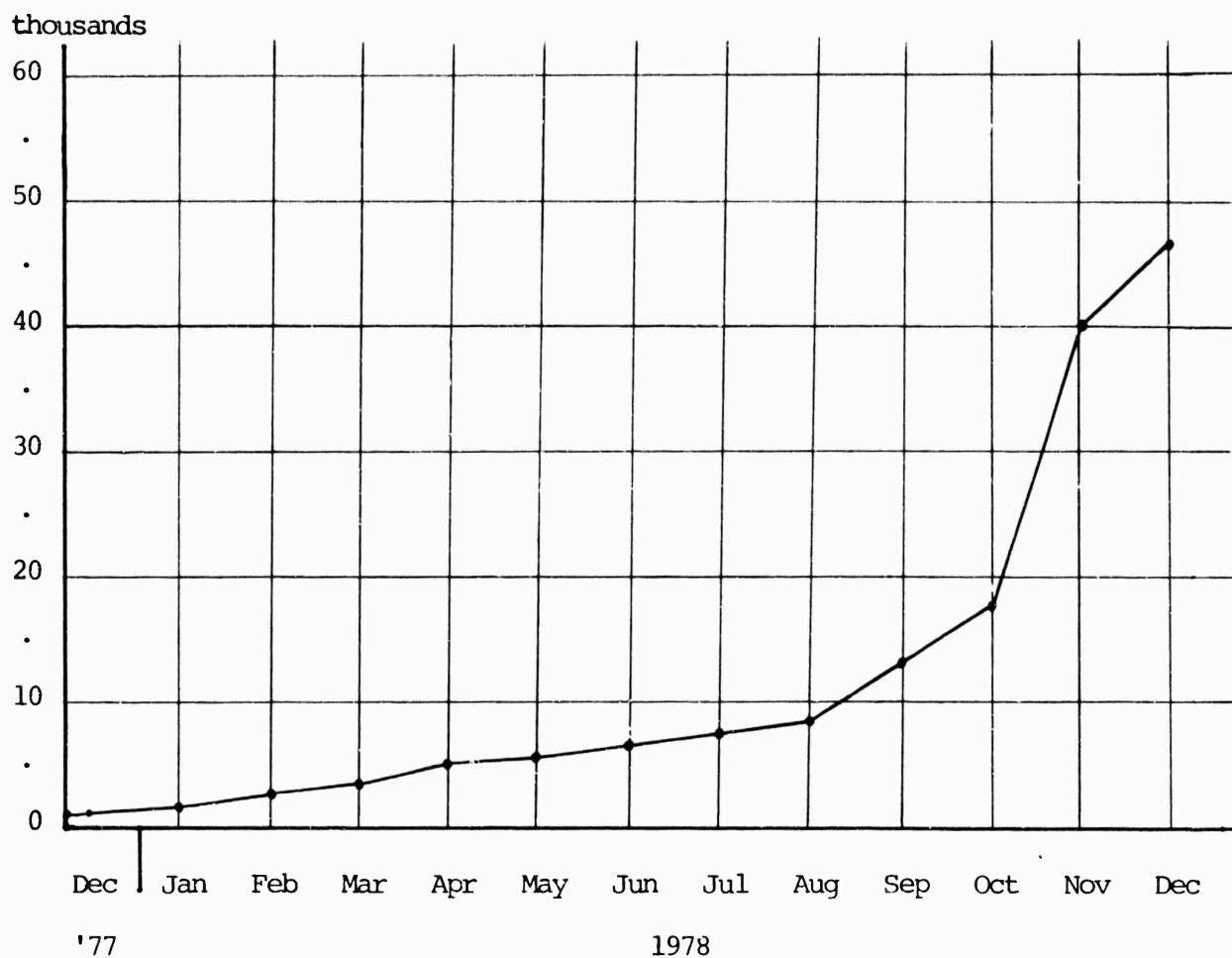
1978

3.   Datacomputer Development

The Datacomputer software is in a primarily operational
state, rather than in a developmental phase. As a result,
modifications are of an evolutionary sort. Below are
described the changes that occurred in 1978 in the
Datacomputer, its documentation, and a special status
server program for the Datacomputer. This section ends
with a brief mention of some future development prospects.

3.1 Version 5

The Version 5 Datacomputer was put into service early in
1978 replacing the Version 4 Datacomputer. There were two
new features of particular significance in Version 5 as
described below.

3.1.1 Message Deadlock Solution

In the past, network communication lockups were found to
occur with the Datacomputer under certain conditions.
They were due to simple user programs that are unable to
listen for Datacomputer error or information messages when
they are transmitting to the Datacomputer or while they
are receiving data from the Datacomputer over a separate
connection. Particularly with the small network buffer
allocations some users provide, it is possible for the
Datacomputer to send enough message text to the user that
the buffers available for that channel fill up. This
results in the Datacomputer waiting for the user program
to read some of this message text and free up some space.
Unfortunately, the user program may be completely intent
on sending or receiving data for which it is in turn
waiting for the Datacomputer. Hence a deadlock.

An elegant and complete solution is for user programs to
have multiple processes so that one can always be
listening for messages from the Datacomputer. (This
technique is used in DCLINK as described in Section
2.4.1.) However, multiple processes may not be supported

on the user's host system or the   user   may   not   wish   to
implement them.   A   solution for such simple one-process
users has been   incorporated   in   Version 5.   It   is   an
optional   mode   in which a limited number of messages that
would normally be sent to a user are inhibited during data
transfers and are given to the user all at once at the end
of the transfer.   Of   course,   error   messages   that   are
sufficiently severe to terminate the transfer are sent and
also   cause any inhibited messages to be sent to the user.

3.1.2   Pre-Compiled Request

In some cases, it is desirable to execute   a   Datacomputer
request   several   times.   (The request may have different
effects depending on the value of various   data   files   or
streams   or its repeated execution may have some desirable
cumulative effect.)   It is more efficient in such   a   case
to   store   the request in a partially compiled form rather
than having to re-input it each time.

The previous version of the Datacomputer provided   such   a
stored request facility only within a single user session.
The   Version 5 Datacomputer   provides   a   more   general
pre-compiled request facility in   which   requests   can   be
stored   in   the Datacomputer between sessions, in a manner

similar to file . The original text is also stored and,
if certain conditions have changed so as to invalidate the
pre-compiled version, the stored request is automatically
re-compiled on execution.

3.2 Version 5 User Manual

For several years, each new major version of the
Datacomputer was accompanied by an update to the previous
complete manual, the Version 1 User Manual. In early 1978
however, a completely new manual was published. This
manual had been in preparation since mid 1977.

The manual includes chapters which cover the following
topics from a user viewpoint:

                    . Datacomputer Facilities and Applications

                    . The Datacomputer Operating Environment

                    . The Directory

                    . Data Security

                    . Data Description

                    . Datacomputer Commands

                    . Elementary Data Manipulation

                    . Efficient Retrieval Techniques

                    . File Groups

. Chaptering and Variable-Length Updating

. Datacomputer Space Allocation

. Interface to the Datacomputer

The Manual also has a thorough index and three appendices
which give a syntax summary of Datalanguage, a list of
Datacomputer reply messages, and charts of Datacomputer
character conversions.

3.3 Datacomputer Status Server

It is frequently useful for remote users to be able to
inquire about the general status of the Datacomputer and
about the status of the network connections to the various
remote jobs using the Datacomputer. The most common cases
are those in which the Datacomputer appears to be
unavailable and those in which the user is connected to
and using the Datacomputer but is concerned over the
status of his request.

To satisfy these needs, an independent status server
program has been implemented. This server runs at
CCA-TENEX but as a separate job independent of the
Datacomputer. When the user connects to this server by
using the DCSTAT program (or Telnetting to socket 141

(decimal) at CCA), a message is sent to the user reflecting most facets of Datacomputer status. This message includes a general status note set by CCA personnel. Since it is presumed that anyone interested in Datacomputer status has some software capable of reading Datacomputer reply messages, the status server formats its status message using the same rules that actual Datacomputer replies obey.

This year, the status server was augmented by adding information to that given for the status of network connections. As well as giving the socket numbers, network, and buffer status of such connections as seen from the Datacomputer, it now also gives the total number of bits that have been transmitted on open connections. This is particularly useful on connections where the receiver is much slower than the transmitter, resulting in buffers filling up and the transmitter being blocked most of the time. Previously it was hard to tell if any progress was being made; now it is possible to see the number of bits transmitted increasing.

A technical bulletin documenting the status server was prepared and distributed with the Version 5 User Manual.

3.4   Version 5/2

Improvements were made in Version 5 during 1978 resulting
in Version 5/2 which, after testing, went into full
operation on 13 December.   These changes consist of a
number of upward compatible additions and one significart
bug fix.

The bug fixed was related to a low probability failure in
interlocking between several users who are trying to
access the same data while the data is being moved from
TBM storage to disk storage.   The problem first made
itself evident on April 1,   1978, when two users were
trying to read the same seismic data. A temporary fix was
used until Version 5/2 was installed.

The several enhancements made in Version 5/2 can be
divided into operational improvements and user facilities.
The user facilities added were requested and funded by the
SDD-1   project   under   a   separate   contract   (No.
N00039-77-C-0074).   The   user facilities included an
addition to the MODIFY command to allow users to change
file names, a new REDESCRIBE command to allow limited
changes to the description of a file without affecting the

data stored in the file, and a new SUBSTRING function to allow the extraction of substrings.

The operational enhancements were those necessary to provide a family of additional operator commands for investigating and repairing problems with parts of the Datacomputer directory. This includes a command for printing limited information on all active directory nodes and commands for printing various types of detailed information on particular directory nodes. If a problem exists and can be solved by backing up to another version of the data for a file, a command is provided for this. In cases where other ad hoc steps are necessary, the commands provide aid in automatically reading in the relevant information, locking it so other users can not interfere, and writing the possibly modified information back afterwards. Before these commands were added, it was almost always necessary to run the Datacomputer in a stand-alone operator only configuration to investigate and fix such problems. Now, when they occur, they can be fixed more rapidly and usually without halting normal service.

3.5 Future Development Prospects

An extensive list is kept of possible improvements in the
Datacomputer system.   Several items on this list related
to improved operation are expected to be given high
priority in 1979. These include moving more Datacomputer
directory information out of cramped TENEX disk space into
Datacomputer controlled storage space,   and   possible
efficiency improvements, especially on large files such as
are frequently encountered in the seismic application.

In   addition,   it   is   planned   to implement direct
Datacomputer access from the National Software Works
system using the MSG interprocess communications protocol.
A significant upgrade to the TENEX operating system under
which the Datacomputer operates will have to be made to
support this facility.

4.   Signal Waveform File Development

In August 1977, the concept of a compacted database of
event-related seismic waveforms was developed jointly by
representatives of ARPA-NMRO, Vela Seismological Center
(VSC), Lincoln Laboratories Applied Seismology Group
(LL-ASG), and the Seismic Data Analysis Center (SDAC).
The data collection would be stored as a   set   of
Datacomputer files, the Signal Waveform Files (SWF), and
would contain extracted seismic readings -- those directly
associated with events in Event Summary Files (ESF).

Both SDAC and CCA will contribute to this SWF database:
SDAC is responsible for storing array data and other data
readings which do not exist in any other form on the
Datacomputer.   Such data will be stored directly into the
Signal Waveform File in segments corresponding to entries
in the Event Summary Files.

CCA is responsible for storing Seismic Research
Observatory (SRO) data segments which already exist on the
Datacomputer, but which are embedded in the very large
basic seismic data files: the non-array long-period files
(NLPF) and the non-array short-period files (NSPF).

Readings containing the data for seismic events which have
been detected and marked in an Event Summary File by SDAC,
are simply copied into the Signal Waveform Files, once
these stretches of basic data have been located and
delimited. Figure 4.1, based on a diagram supplied by
SDAC [TELEDYNE], depicts the overall data flow through the
system and gives an indication of the order of processing.

In 1978, CCA designed a program [EASTLAKE & SATTLEY a] to
accomplish its part of the project. The SWF-D ("D" for
Demon) program was implemented under this contract.
Figure 4.2 shows the data of concern to SWF-D. The work
has been described in detail in previous technical reports
submitted by CCA [SATTLEY a, SATTLEY b]; a general
description is given below in Section 4.1.

CCA also participated in the general planning of the SWF
project [EASTLAKE & SATTLEY b]. Representatives from CCA
attended meetings at SDAC on 16 January and 13 April to
discuss and plan the SWF. We also wrote two memoranda,
which were relevant, as follows: "Comments on the SDAC
Mass Store Retrieval Guide" (17 January) and "TBM Drive
Allocation and the SWF" (20 March). Later in the year,
two changes were made in the Event Summary File
description, at CCA's suggestion, for the SWF project.
They were the addition of a virtual index and the
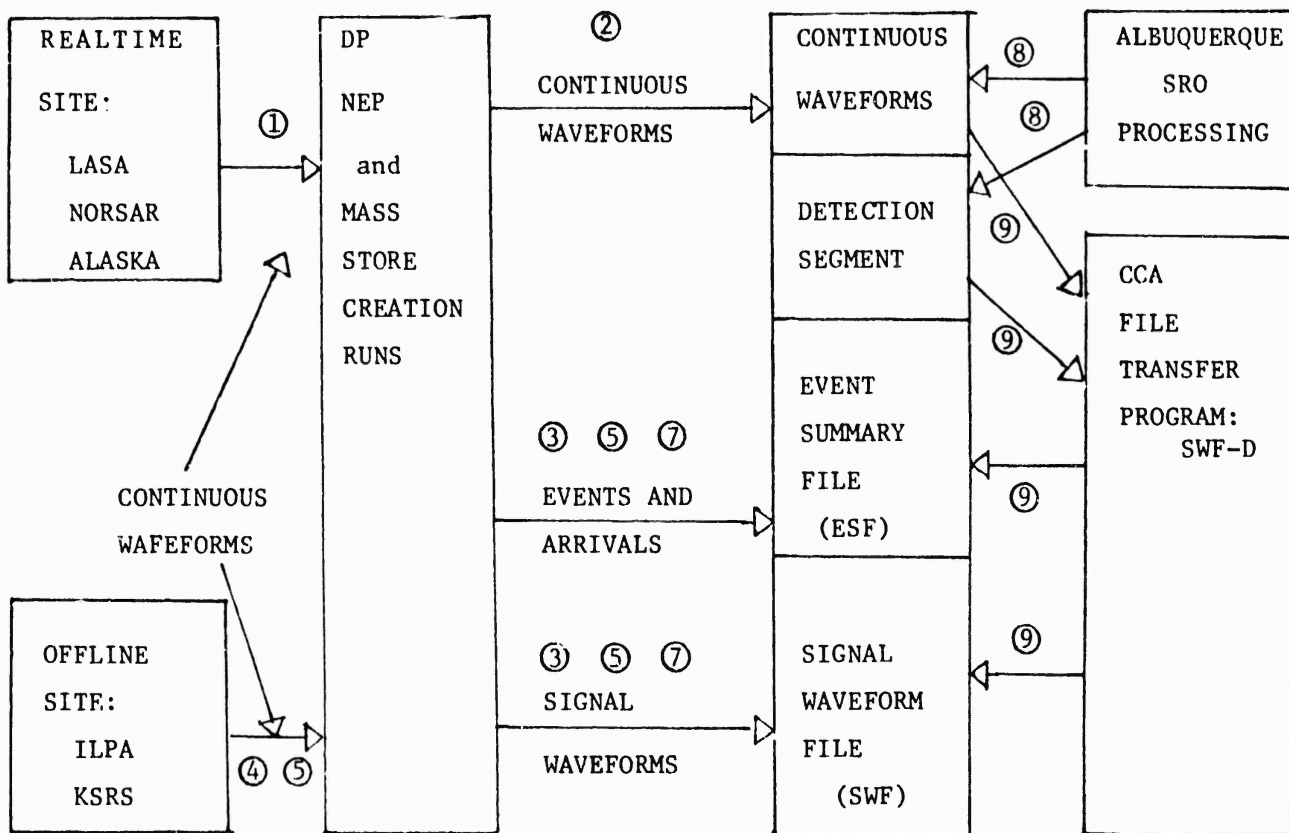
1 - Waveform data from realtime sites arrives at SDAC

2 - Some data from realtime site goes to Datacomputer

3 - ESF/SWF including realtime site signal waveforms go to Datacomputer

4 - ILPA field tape arrives at SDAC

5 - ILPA signal waveforms go to Datacomputer

6 - KSRS field tape arrives at SDAC

7 - KSRS signal waveforms go to Datacomputer

8 - Albuquerque reviewed SRO data goes to Datacomputer

9 - CCA FILE TRANSFER PROGRAM completes signal waveforms
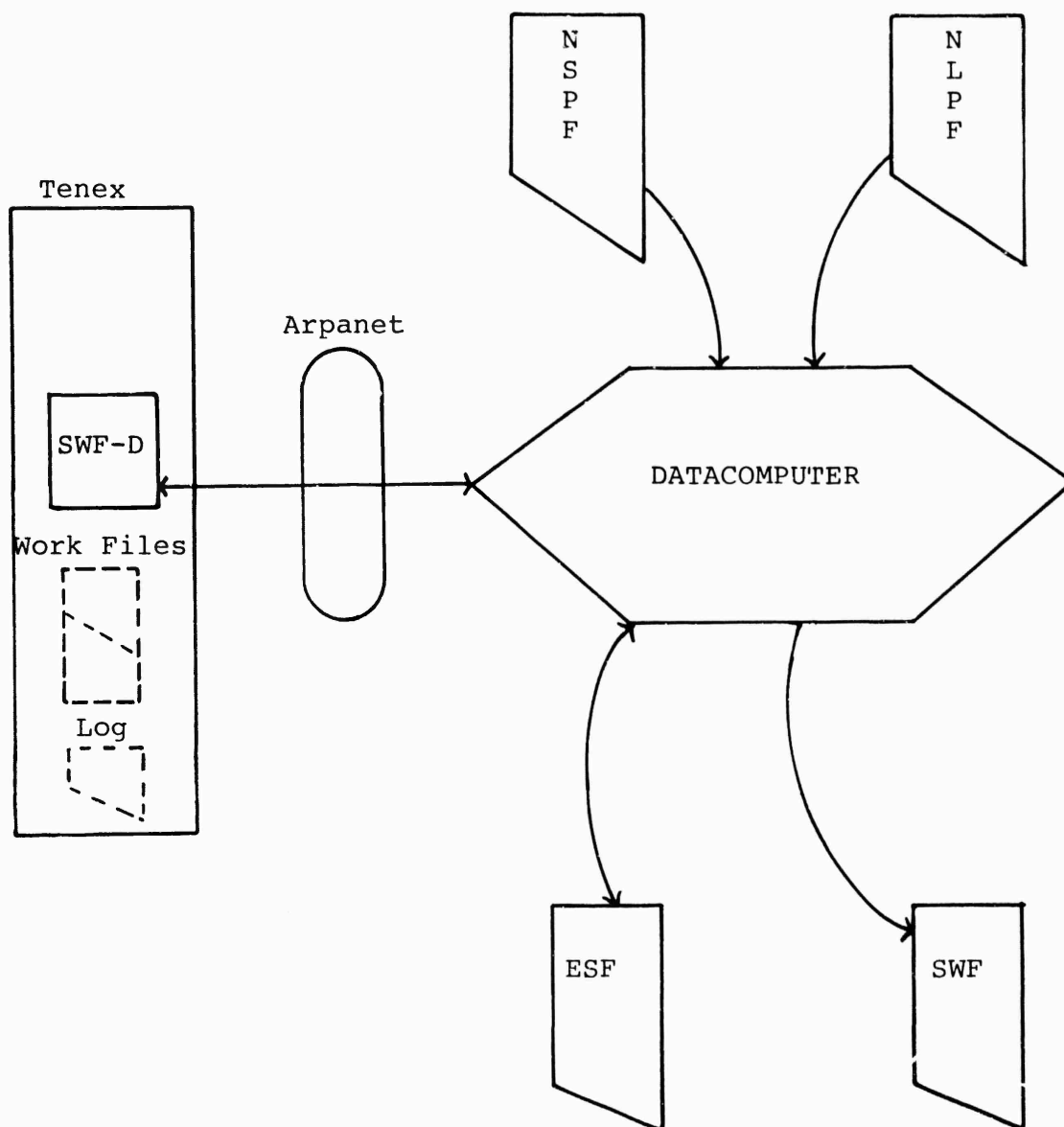
elimination of two unnecessary inversions that made it
impossible to update certain fields.

4.1 Tasks of SWF-D

The SRO seismic data, from which SWF-D must extract
portions, is stored into the Datacomputer from magnetic
tape by the Albuquerque Seismological Laboratory. The
data arrives at the Datacompter several months after real
time.   Long period SRO data is recorded and later
transmitted to the Datacomputer continuously. The higher
bandwidth short period SRO data is only recorded and later
transmitted to the Datacomputer when it exceeds a
threshold set at the seismic instrument. Because of this
data routing, some of the SWF candidates may not yet be on
file at the time they are requested, or they may be only
partially available.

For the long-period files, expected data availability is
based upon the date(s) specified in the periodic report of
"Digital Seismic Data Transmitted to Datacomputer" via
Arpanet message from the Albuquerque Seismological
Laboratory (ASL).

For short-period files, determining data availability is a bit more complicated. To help in checking for which readings may be available, the SWF-D program pre-processes the short-period files and generates a list of the recorded detections on each full cycle through its task queue. The list, sorted into monthly Datacomputer files, the SPDET files, is used as a map which can be consulted for deciding whether to attempt to retrieve the desired data.

The SWF-D program comprises separately-compilable components -- modules -- which operate independently of one another to accomplish a set of prescribed tasks:

- The ESF-checking Module, to accumulate requests for moving waveform segments by sifting the Event Summary Files for flagged arrivals;
- The Waveform-copying Module, to copy the desired long-period and short-period waveforms into the Signal Waveform Files; and
- The SPDET-mapping Module, to generate a map of the short-period detections for use when determining expected data availability.

Each program module is capable of recognizing when there is work for it to do, logging in a separate Datacomputer job, generating a script file for the session, interacting with the Datacomputer to manipulate the relevant files, reporting task progress on an operations log file, and terminating the session when a pre-programmed quota of work has been performed.

Because of the program's modularity, each task-handler could be tested individually and, when running error-free, could be scheduled for regular operation in detached background mode. The ESF-checking and SPDET-mapping Modules have already been placed in service; they have worked their way through the set of NSPF and PESF files beginning with the data for 1 January 1978, and have recently reached 25 July 1978. The Waveform-copying Module has been operated in debugging mode only and will need further testing for bad-data protection and for correct handling of massive data.

4.2 SWF Usage

The SPDET files are already being used in studies
conducted by the Lincoln Laboratory Applied Seismology
Group [LINCOLN].

It is anticipated that the attainment of full service
operation in 1979 will result in an increase in
Datacomputer service to the seismic research community.

5. The SIP

The SIP is a dedicated minicomputer communications system developed and operated by Computer Corporation of America. It interfaces real time seismic information from the Velanet to the Datacomputer [DORIN & EASTLAKE].

Below we give a general description of the SIP and the changes in the SIP and the seismic data flow in 1978.

5.1 General Description

A primary function of the world wide Vela Seismological Network (Velanet) is the collection of real time seismic data from arrays of seismometers. This data is sent over leased lines and the Arpanet to the Communications and Control Processor (CCP) at the Seismic Data Analysis Center (SDAC) in Alexandria, Virginia. From the CCP this data is immediately distributed to various processors and to the Datacomputer, via the SIP, for storage as shown in Figure 5.1. This figure also shows the SDAC and Lincoln Laboratories UNIX systems which are presently the primary
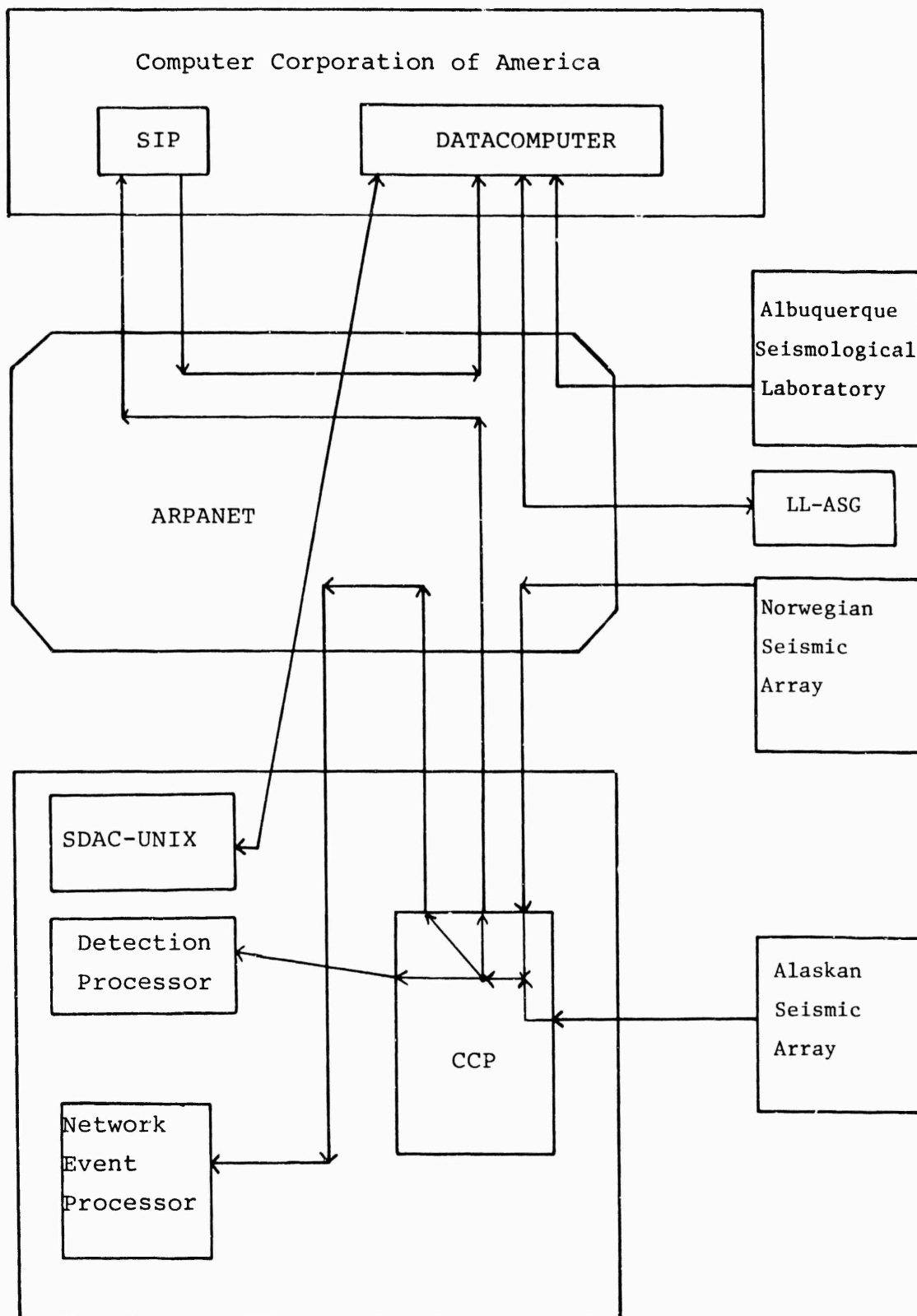
systems used by researchers retrieving seismic data from
the Datacomputer.

The components of the Velanet that handle real time
seismic array data, except for the Datacomputer, are
dedicated systems designed to receive data in real time.
The Datacomputer, however, operates within a non-real-time
operating system and serves the general Arpanet community
as well as the Velanet. Furthermore, it is occasionally
unavailable due to scheduled and unscheduled maintenance
work. To isolate the Datacomputer from these real time
requirements, the SIP was implemented to receive real time
data from the CCP, buffer and reformat this data on disk,
and periodically forward the data to the Datacomputer.

The SIP's hardware structure is shown in Figure 5.2. It
is implemented on a DEC PDP-11/40 computer. The SIP has
an Arpanet interface, two RP-04 disk drives for buffering
data, an operator's terminal, and a status display screen.
With the present bandwidth and disk structuring, about two
days of data can be held by the SIP.

Besides processing seismic data, the SIP software provides
for operator communications between itself and the CCP.
It also sends messages to the CCP for each chunk of data
when the data has been properly filed in the Datacomputer.
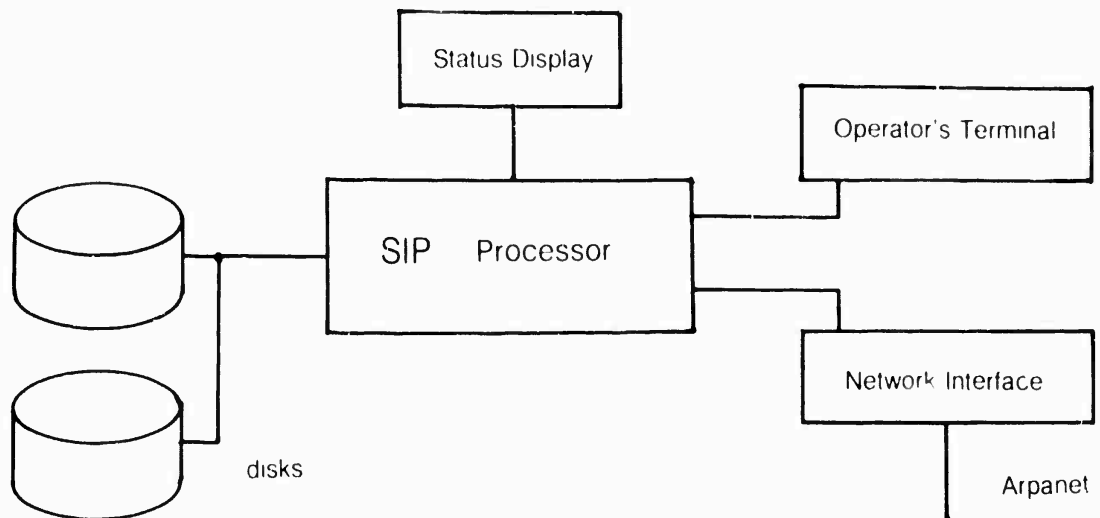
5.2 SIP Development in 1978

As set up at the end of 1977, the SIP received data from
three seismic arrays: LASA, the   large   aperture seismic
array in Montana; NORSAR, the Norwegian array; and ALK,
the Alaskan array. Each of these   arrays   provided both
long period (one sample a second) and short period (ten or
twenty samples a second) seismic data.

Since it was expected that the number of arrays would increase, the SIP's directory structure was modified. It had originally allowed for up to 4 sites. The modifications that were initially implemented provided for up to 8 sites and modifications for higher numbers were investigated. Further details are given in CCA Technical Report 79-08 [Eastlake b].

However, the expected increase in the number of sites has not occurred. Instead, the present direction of development for the seismic system is towards the SWF or Signal Waveform Files. These files are intended to include all the "interesting" planet data and are described in Section 4 above.

The changes that actually occurred during 1978 were a cessation of real time short period data being sent to the SIP, which required a very minor modification to the SIP, and later the termination of data from the LASA site when that site was permanently decommissioned.

The reliability of the SIP was demonstrated during the Great Blizzard of 1978. Despite repeated local power failures, the SIP successfully restarted itself whenever power was restored.

We continue to receive and process real time long period
data through the SIP, and the SIP directory system
modifications are still in place to enable us to easily
adapt to data from an increased number of seismic arrays
when the need arises.

6.   The TBM Mass Storage System

It is the TBM Mass Storage System that enables the
Datacomputer to store and access such a large volume of
on-line data. The TBM is described below. This
description is followed by a section on some 1978
enhancements to a TBM utility program and a section on TBM
reliability.

6.1 General Description

The CCA Datacomputer is equipped with the first public
installation of the Ampex Tera-Bit Memory (TBM) System
[Eastlake a]. This device uses video tape technology to
achieve a maximum on-line capacity of 3 trillion bits on
64 tapes. Maximum seek time to any bit is 45 seconds.
Maximum data transfer rate is 5.3 million bits per second.

The TBM at CCA is equipped with two dual tape transport
modules so at most four tapes, or 175 billion bits
(equivalent to 220 IBM 3330 type disk packs) can be
available on-line. All equipment between the transports

and the Datacomputer central processor is non-redundant in
the CCA configuration as shown in Figure 6.1.   There is
one transpor driver (necessary for a tape to be in
motion), one data channel (necessary to encode and decode
digital information to and from the broadband analog
signal on tape), one system control processor to
coordinate and direct the other units, and one channel
interface unit that connects the TBM system to the
Datacomputer's PDP-10 system.

6.2   TBMUTL Enhancements

In order to maintain the TBM, it is necessary to test its
operation from the CCA PDP-10 host computer.   This is
normally done daily after the early morning TBM
maintenance period. TBMUTL (for TBM UTiLity) was written
to answer this need.   It provides facilities ranging from
the ability to execute all the individual TBM device
operations available through the modified CCA TENEX system
to a convenient way of executing a general canned test on
all TBM drives sequentially.

TBMUTL is relatively mature and only two minor
enhancements were made to it this year. First, some
operations were added to the most commonly used canned

to/from Datacomputer via SA-10   Operator Console
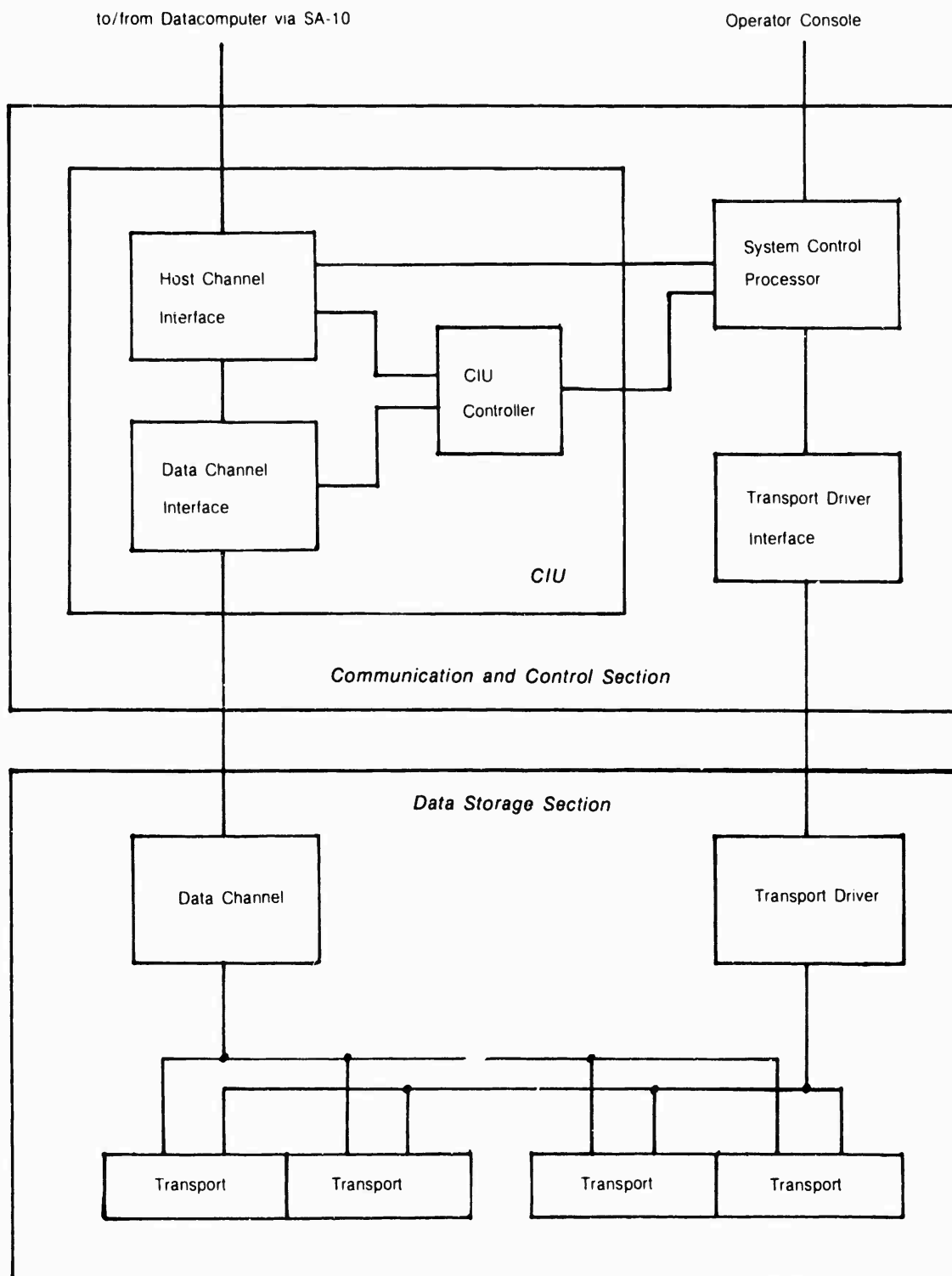
test so as to more throughly check TBM operations. One of
these new operations deliberately caused a minor error to
check the proper operation of most of the error detection
and processing logic. Secondly, the automatic sequencing
through drives for the predefined canned test sequences
was made more general. Previously it was possible only to
specify all drives or any single drive for a test. As a
result, if a drive was unavailable, it was necessary to
have a person in attendance to sequence the test through
the drives skipping the bad one. It is now possible to
specify any set of drives for the test to sequence
through.

6.3 TBM Reliability

As shown in the Table 2.3, the TBM is the least reliable
of the components making up the Datacomputer. The reason
for this is that the TBM installation at CCA is, except
for the four tape transports, entirely non-redundant.

An unfortunate aspect of the TBM periods of unavailability
is that they are relatively long. It is not a matter of a
system crash or the like, where service can be restored
within, say, an hour. Rather it is an occasional obscure
failure deep within the controller hardware that may not

Datacomputer TBM Prime Time Unavailability

Jul - Dec 1977   - 172.9 hours
Jan - Jun 1978   - 82.0 hours
Jul - Dec 1978   - 43.0 hours

Breakdown by Hardware Module Type

	JulDec77	JanJun78	JulDec78	Overall
C.I.U.	47.5	29.3	11.0	37.2
Data Chan.	21.3	22.0	-	18.6
TransDriver	10.6	-	-	6.2
Drives	7.9	16.8	85.5	21.3
S.C.P.	-	-	-	-
Undiagnosed	12.7	31.9	3.5	16.7
	100.0	100.0	100.0	100.0

----

be fixed for many hours or, in some cases, a couple of days.

Table 6.1 gives a detailed breakdown of reasons for TBM non-availability from mid-1977 through the end of 1978. Note that the hours of prime time unavailability due to the TBM have been declining markedly.

Most TBM problems were associated with controller modules of the TBM except for an unprecedented drive related problem in the second half of 1978. Specifically, in July 1978, our only incident of significant damage to TBM tape with user data on it occurred. The unavailability time

PRESENT CCA TBM
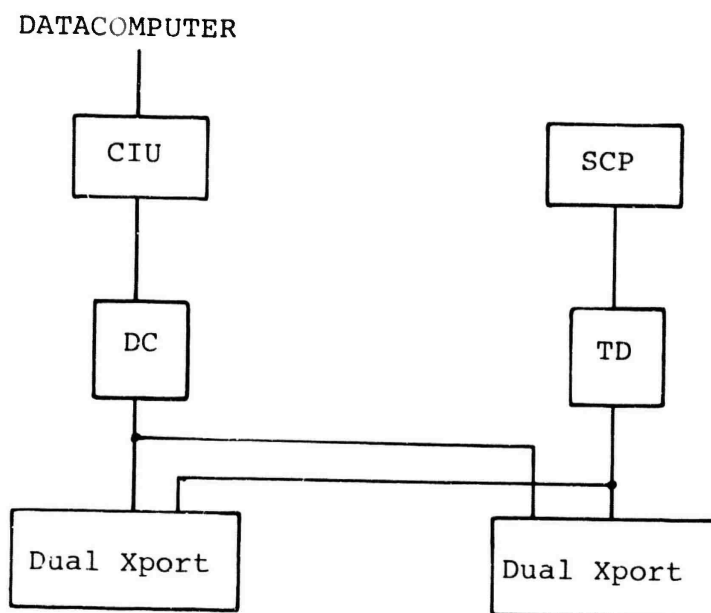
POSSIBLE REDUNDANT TBM
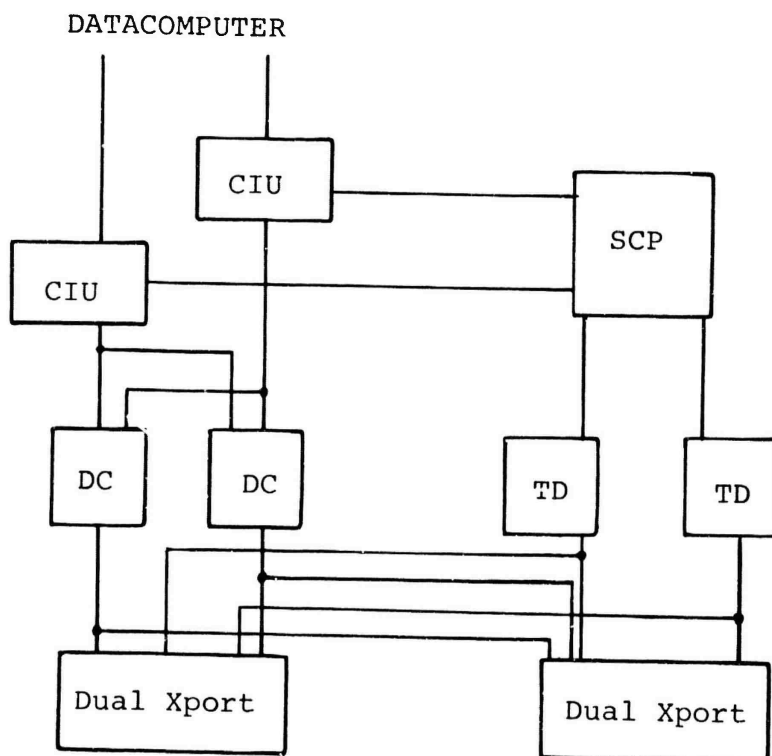
indicated includes the time it took to fix everything up using the backup data copies that the Datacomputer maintains.

Figure 6.2 shows a simplified block diagram of the present CCA TBM installation and of a possible redundant version. All more recent TBM installation., such as the one at the National Center for Atmospheric Research in Boulder Colorado, are more redundant than the CCA installation and have much better availability statistics.

7. The CCA TENEX System

The Datacomputer runs as a single, although unusually
large and complex, user job on the CCA-TENEX host computer
under the TENEX operating system. Many modifications had
to be made in TENEX to accommodate the special
requirements of the Datacomputer job and the special
hardware in use on the CCA Datacomputer system. Figure
7.1 shows the hardware structure of the system.

The general modifications and additions that had to be
made to CCA-TENEX include both changes to the operating
system and additional utility programs that are not part
of the Datacomputer proper. These modifications and
additions include the following:

1. improved efficiency in the network interface code
    for the high volume file-transfer-like data streams
    sent and received by the Datacomputer;

2. device code was implemented for using the Ampex TBM
    Mass Storage system and CalComp 3330-type disks;

CORE MEMORY . . . . CORE MEMORY   CORE MEMORY

MEMORY BUSSES

SA-10   DF-10   BBN PAPER

                          I/O BUSS   PDP-10
                                                CPU

DATACOMPUTER   TENEX
DISKS   DISKS

                                                LINE
                                                PRINTER

TBM

                                                LOCAL
                    PAPER   TERMINAL
                    TAPE   CONTROLLER

                    MAGNETIC   ARPANET
                    TAPE   INTERFACE
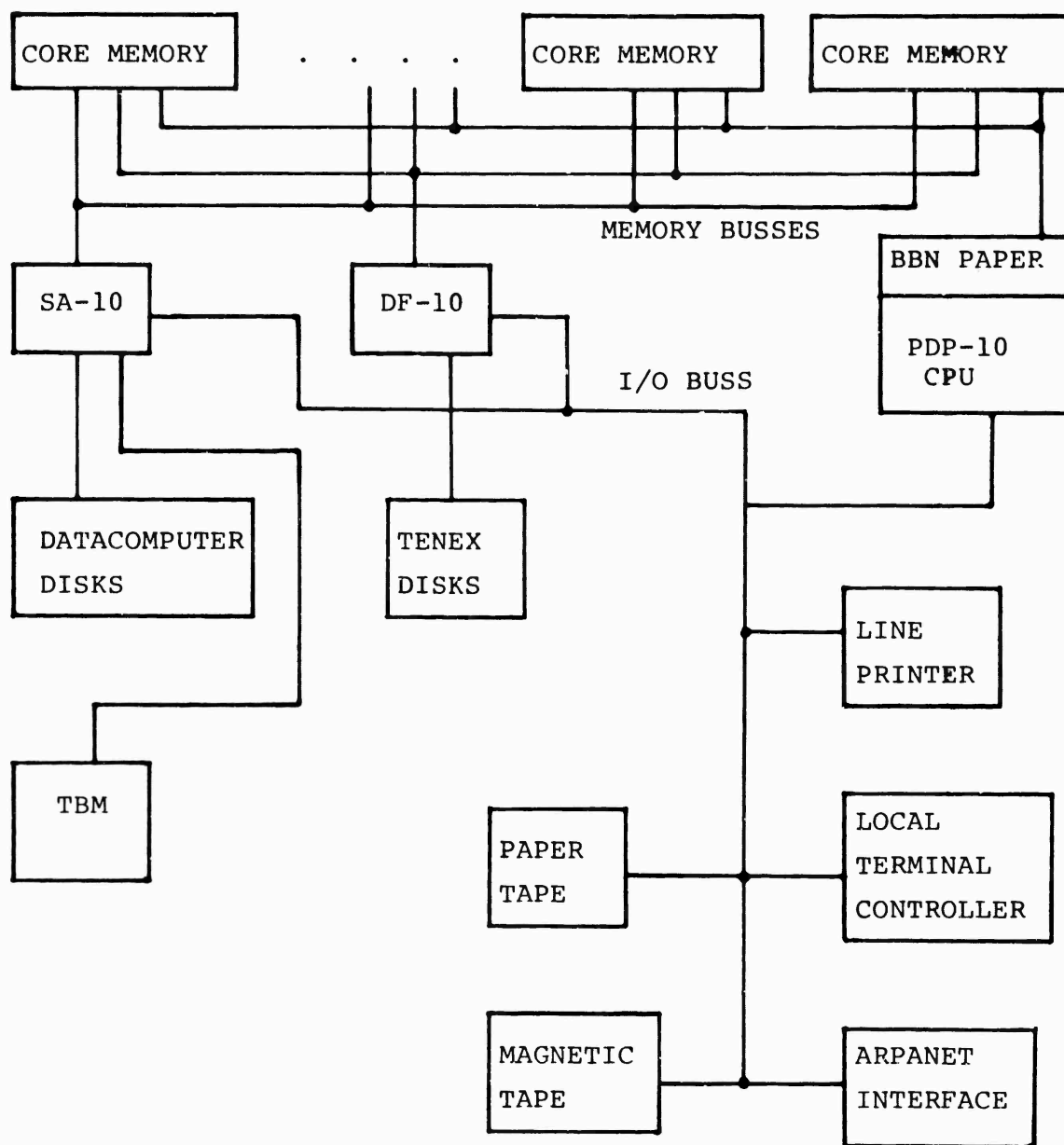
3. additional statistics gathering code was added to
    aid optimization and analysis of operating system
    performance;

4. a separate network server program was augmented to
    provide status output on the Datacomputer [Eastlake
    c];

5. a utility program called TBMUTL was written to
    assist in TBM maintenance;

6. a utility program was implemented called ALL that
    runs as a background job in CCA-TENEX and keeps an
    eye on various system resources; and

7. other utility programs were added including one for
    cleaning up old audit trail files of user sessions
    in the Datacomputer's TENEX directory.

Enhancements made and problems encountered in 1978 are
described in the following sections.

7.1 Datacomputer Related Changes

Three Datacomputer related changes were made in the
CCA-TENEX system in 1978. One involved a change to the
background program ALL, one related to handling certain
errors indicating that the TBM was unavailable, and one
involved expansion of available disk space.

To maximize Datacomputer availability, it is desirable to
bring hardware failures affecting the Datacomputer to the
attention of someone who can rectify the situation. To
this end, the ability of the ALL program (named after the
cry of the watch: "All's Well") was enhanced. In the
past, in addition to appending periodic statistical output
to several log files maintained by CCA-TENEX, ALL
monitored the load average and amount of free TENEX disk
space. If the system is excessively loaded or has a
dangerously low amount of disk space available, a warning
is periodically printed out on all terminals. Now ALL
also checks for problems with a TBM drive, with the TBM or
staging disk subsystems, or with the CCA-TENEX network
connection. For problems with any of these, a warning
message is similarly broadcast to alert CCA personnel and
minimize the time to start corrective action.

As mentioned above, one of the items checked by ALL is a flag that indicates that the TBM subsystem is not available. Some very high level errors that were associated with the hardware common to the TBM and staging disk subsystems did not set this flag. Additional code was inserted so that if such errors were occurring, the TBM subsystem would be marked as unavailable.

During the early parts of 1978, TENEX disk space was becoming increasingly tight due to the growth in the Datacomputer directory, audit trail, logging, and accounting files. Although in the longer run it is hoped to move much of this information to other storage than TENEX disk space, the solution adopted in the interim was to add a 7th RP02 disk to the system. This required reassembly of the system and, because the TENEX system has no explicit notion of separate physical volumes, a complete dump and reload of all information stored in TENEX disks at the time.

7.2 Other Changes

A new version of the File Transfer Protocol user program
was installed at the begining of July, 1978.

In late July, new versions of BCPL and BDDT were
installed.   BCPL is a high level language that has
become the preferred tool in developing ...   -10
software systems at CCA.   BDDT is a debugger especially
designed for debugging programs writt in BCPL.

For use in conjunction with BCPL, which distinguishes
upper and lower case letters, and general documentation, a
program called XLA180 was developed near the end of 1978.
This program allowed more effective use of our LA-180 low
speed printed by properly adjusting output to the printers
timing characteristics and thus avoiding data loss due to
printer buffers overflow.

7.3   CCA-TENEX Network Problems

We had some problems related to our Arpanet connection
during 1978 as shown in the unavailability Table 2.3.
Almost all of these can be traced to work on the PLURIBUS
IMP at CCA. This IMP is our local Arpanet communications
node and all network communications must go through it.
Although it is a multi-processor IMP and was in fact
expanded from two to three processors during the year, it
has only one bus. Consequently any hardware work on the
IMP end of any host interfaced to it requires suspension
of service to all the hosts directly connected to it.

In one case, the CCA host interface was unintentionally
modified while work was being performed on the IMP to
install another host interface; it was several days before
the situation was properly diagnosed and repaired.

A.   General Description of the Datacomputer

This appendix contains a brief general description of the
structure of the Datacomputer which is intended to provide
context for other parts of this report.   Persons already
familiar with the Datacomputer may skip over it. On the
other hand, persons desiring an even more detailed
internal description than presented here are referred to
the final Semi-Annual Technical Report for the
Datacomputer Project, contact number MDA903-74-C-0225,
covering July through December 1976, and to the following
two papers, which are available from CCA: Tertiary Memory
Access and Performance in the Datacomputer (CCA-77-11),
and Integrity and Recovery in the Datacomputer, a Machine
for Very Large Databases (CCA-78-06).

It is intended that the direct, immediate user of the
Datacomputer be a program running on an Arpanet host
remote from the Datacomputer. This program may be acting
as an intermediary for a person at a console, or it may be
an automatically started background task that normally
operates without human   ervention, or it might be some
other type of program. See Figure A.1.

Datacomputer and SIP FTR
General Description of the Datacomputer

Page -65-
Appendix A

Datacomputer's User View

Figure A.1

Datacomputer User's View
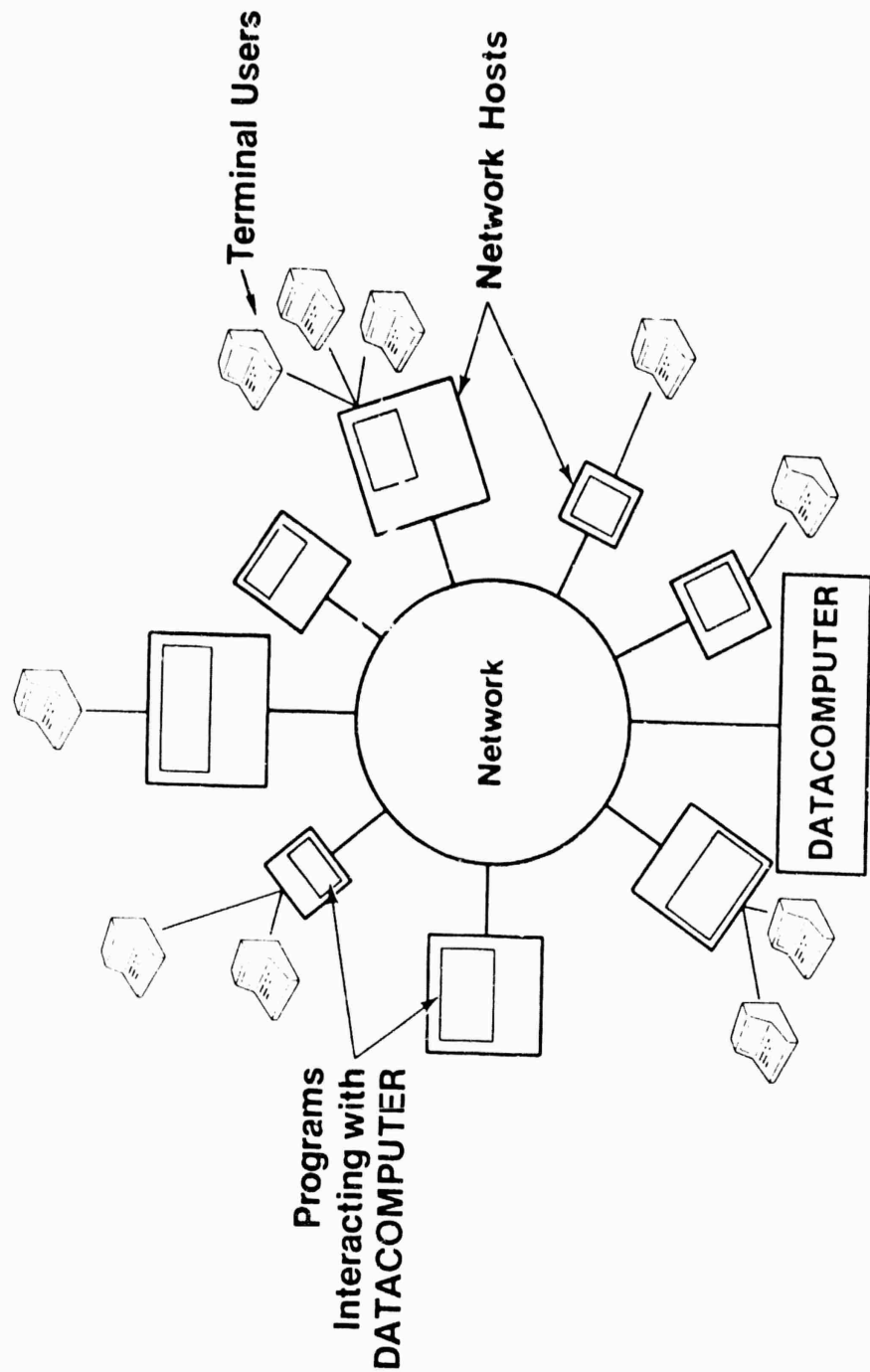
Such a remote program calls the Datacomputer over the
network and establishes a pair of standard uni-directional
8-bit byte network connections. The program then proceeds
to send Datalanguage over one connection while the
Datacomputer replies on the other. Datalanguage is a
uniform high level language which gives the user a
hierarchically structured view of his data. It is
described in the Datacomputer Version 5 User Manual which
is available from CCA.

Actually, the Datacomputer sends a "reply" to prompt the
user program whenever the Datacom.uter is ready to accept
another line of Datalanguage. Similarly, the Datacomputer
keeps the user program abreast of the progress of its
requests with various synchronization, error, and success
messages and comments. All replies have a fixed format
which is designed for easy parsing by the user program.
The reply begins with a number quickly identifying certain
important messages. This is followed by the date, time,
and an internal Datacomputer identification of the message
source. The remainder of a reply provides a
human-readable version of the message to be conveyed.

In the case of serious errors, to assure resynchronization
with the user progi ', the Datacomputer enters a special
mode. In this resynchronization mode, it rejects all

messages from the user, giving an appropriate reply each time, until the user program sends a special message to clear the condition. This mechanism allows the user program to assure that commands sent after an unexpected error are not misinterpreted or partially ignored and partially executed.

Data as well as commands and replies can be transmitted between the user program and the Datacomputer over these initial 8-bit connections but certain characters are prohibited and the connections are fixed at their 8-bit bytesize. More general data transfers can be performed by making use of separate data connections.

The requests, replies, and data for a particular user program are handled by the half of the Datacomputer known as RH or the Request Handler. RH handles the parsing of the user commands, which are in Datalanguage, and the synthesis of replies. For more complex commands, RH takes the user's requests and the data descriptions stored in the Datacomputer and compiles them, in several stages, into code to execute the request.

Data descriptions in the Datacomputer are associated with each file of data. Data descriptions are also provided for data streams to be received from or transmitted into the Arpanet. These streams are known as ports.

Descriptions are set by the user when a file or port is
created.   Datacomputer data descriptions provide for
hierarchical arrangements of structures of diverse data
elements and repetitive lists. Many data types and data
formatting alternatives are provided. This is important
in ensuring that the Datacomputer can communicate
efficiently with its diverse class of remote user
computers.

RH accomplishes most of its activities by calling on
routines in the other half of the Datacomputer, known as
SV or services. SV is a pseudo-operating system in which
RH runs. It is SV that actually has custody of the
Datacomputer's multilevel hierarchical directory tree and
enforces the extensive protection mechanisms of the
system. SV is the part of the Datacomputer that
ultimately communicates with storage devices and retrieves
or stores bits. It views data as a relatively simple
collection of areas subdivided into fixed size "pages" of
data. All of the more sophisticated data description
mechanism is in RH.

A special subpart of SV, called SDAX, moves active data
from slower tertiary memory to faster secondary disk
storage and moves it back when secondary storage is
crowded. The CCA Datacomputer uses an Ampex TBM,

Datacomputer and SIP FTR   Page -69-
General Description of the Datacomputer   Appendix A

Software View of the Datacomputer   Figure A.2
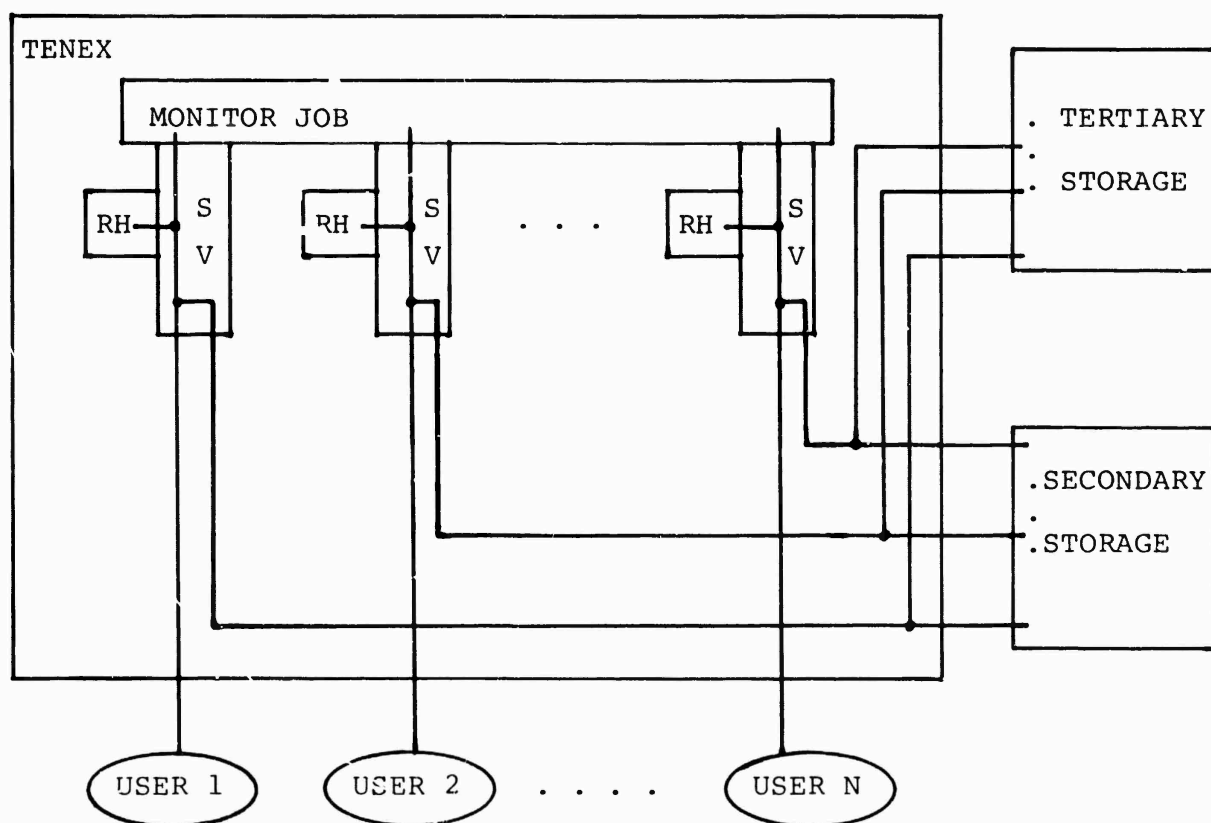
described in Section 6 above, for tertiary storage. SDAX
keeps track of the location of the various copies of each
file data page. It also ensures data consistency by
preventing updates that are not successfully completed
from affecting the original file. SDAX allows multiple
readers and a single updater of a file among the
Datacomputer subjobs. Each reader sees a consistent
version of the file including only updates that were
complete when they opened the file.

At any one time, there are multiple copies of the RH-SV
pair serving users over the network. This actually means
multiple copies of the joint variable area only, because
they are implemented using reentrant code. The RH-SV
pairs are called subjobs as they are all part of one job,
under a monitor process, running in our modified TENEX
operating system as shown in Figure A.2.

B.   Developmental History

The developmental histories of the Datacomputer and SIP
are reviewed in this appendix so as to put our current
work in perspective. Persons familiar with this history
may skip this material.

B.1 Datacomputer Development

Figures   B.1 and B.2 give a quick overview of the
contractual and developmental history of the Datacomputer.
Since only the most important event of a quarter is given,
a similar event may be overshadowed in a different
quarter.

The initial "Network Data Handling System" contract
starting in 1971 funded the initial conceptualization,
design, and demonstration of the system. The installation
of the PDP-10 and large secondary storage was included.
The system developed in 1972 and 1973 used only disk
storage but otherwise provided the same framework for
shared remote use, on a simplified basis, as the present
Datacomputer.

Datacomputer Development 1971-75

Year	Qtr.	"Network Data Handling System"
1971	1	Concepts & Fundamental System Design
71	2	Unicon Interface Designed
71	3	Various Working Papers
71	4	PDP-10 Delivered
1972	1	PDP-10 Accepted
72	2	TIP Installed
72	3	
72	4	Various Working Papers
1973	1	ICCC Demo
73	2	DMCG 1st User
73	3	O/9 Release
73	4	O/9 Manual
1974	1	O/9.7 Up
74	2	O/10 Release
74	3	O/11 Work
74	4	O/11 Up
1975	1	V.1 Work
75	2	V.1&XD Checkout
75	3	V.1 Release
75	4	File Groups Design

"Datacomputer Support of Seismic Data Activity"

Contact Seismic Community
SIP/Site Specified
TBM Ordered
SIP Installed
SIP on Net
IMP/TIP Swap
TBM Site Work

"Datacomputer"

Datacomputer Development 1976-80

Year	Qtr.	"Datacomputer"	"Datacomputer Support of Seismic Data Activity"
1976	1	SDAX Testing	TBM Installed
76	2	TBM Interface	SIP Operational
76	3	File Groups Work	TBM Acceptance Test
76	4	V.2 Up with TBM	
1977	1	V.3 Up	
77	2	New CCP-SIP Protocol, Final TBM Acceptance	
77	3	V.4 Up	
77	4	V.5 Up	
1978	1	SWF Planning	
78	2	V.5 Manual Complete	
78	3	V.5 Manual Distributed	
78	4		"Enhancement of Datamodule 1"
1979	1		
79	2		
79	3	(Proposed)	
79	4		
1980	1		
80	2		
80	3		
80	4		

"Datacomputer & SIP Operations"

From 1974 through 1976, CCA held two contracts. The
"Datacomputer" contract financed the transformation of the
prototype framework Datacomputer into a highly useful data
utility. Facilities of all sorts were added during these
years including file security, tertiary memory support,
accounting, expanded data description facilities including
variable length containers and virtual containers,
expanded data manipulation facilities including
conditionals, blocks of requests with local variables,
and, lastly, a number of critical efficiency improvements.

Meanwhile, the "Datacomputer Support of Seismic Data
Activity" contract financed the SIP and TBM as described
in the section below.

1977 was the first year of the combined "Datacomputer and
SIP Operations" contracts. Primary concentration in 1977
was on operational improvements and user service. Many
new users were contacted some of whom have become regular
users. Final acceptance of the TBM did not occur until
the second quarter of 1977 when various software changes
within the TBM were delivered and verified.

B.2 SIP and TBM Development

The hardware configuration of the SIP was specified around
the middle of 1974 at the time that the site modification
at CCA to accommodate the SIP and the mass storage system
was specified and the Ampex TBM mass storage system was
ordered.

The SIP hardware was delivered in early 1975 and the SIP
was connected to the Arpanet in June of 1975. The volume
of seismic data to pass through the SIP and problems
encountered in early tests indicated the necessity to
upgrade the local Arpanet node at CCA. This was done in
late 1975 when a model 316 TIP node was replaced with a
model 516 IMP.

The SIP was fully operational and hundreds of hours of
seismic data had been retransmitted to small test files in
the Datacomputer by mid-1976. Further work was done in
1976 on improving the SIP's robustness in the face of
hardware and communications troubles. Continuing problems
with CCP<->SIP communications led to the design of a new
communications protocol to be used on this link.
Meanwhile, the TBM mass memory was accepted and integrated

with the Datacomputer and on October 1, 1976, the SIP-Datacomputer system became fully operational storing data into the mass storage system.

In the early parts of 1977, the new CCP<->SIP protocol [BBN] was implemented and the CCA local Arpanet node was further upgraded to a PLURIBUS IMP. These st   cleared up the previously chronic communications probi

Since then, the SIP has been in virtually continuous operation, successfully receiving, buffering, reformatting, and forwarding seismic data.

References

[BBN]
        BBN   Report   No.   3460,   "VELANET Communication
        Protocols between the CCP and the SIP and between
        the CCP and the DP", 5 April 1977.

[CCA]
        Computer   Corporation   of   America,   Datacomputer
        Version 5 User Manual,   Cambridge,   Massachusetts,
        July 1978.

[DORIN & EASTLAKE]
        R.   H.   Dorin and Donald E. Eastlake, III; "Use of
        the   Datacomputer   in   the   Vela   Seismological
        Network"; Technical Report No. CCA-77-08, Computer
        Corporation   of   America,   575   Technology Square,
        Cambridge, Massachusetts, 02139, April 15, 1977.

[DORIN & SATTLEY]
        Dorin,   R.H.   and   Sattley,   J.Z.   Very Large
        Databases: Final Technical   Report,   Technical
        Report   No.   CCA-77-10,   Computer   Corporation   of
        America,   575   Technology   Square,   Cambridge,
        Massachusetts, 02139, August 30, 1977.

[EASTLAKE a]
        Eastlake,   III,   D.E.   "Tertiary Memory Access and
        Performance in the Datacomputer", Technical Report
        No. CCA-77-11, Computer   Corporation   of   America,
        575 Technology Square, Cambridge, Massachusetts,
        02139, June 30, 1977.

[EASTLAKE b]
        Eastlake,   III,   D.E.,   "SIP   Accomodation   of
        Additional Seismic Data Sites", 29 January 1979,
        Technical   Report   CCA-79-08, Computer Corporation
        of   America,   575   Technology Square,   Cambridge,
        Massachusetts 02139.

[EASTLAKE & SATTLEY a]
        Eastlake,   III,   D.E.,   and Sattley, J.Z., "Program
        Design   for   SWF-D: The   Signal   Waveform   File
        Demon", 30 June 1978, Technical Report CCA-78-10A,
        Computer   Corporation   of   America, 575 Technology
        Square, Cambridge, Massachusetts 02139.

[EASTLAKE & SATTLEY b]

Eastlake, D.E., III and Sattley, J.Z., "Assistence in the Design of the SWF System", January 30, 1979, CCA-79-13, Computer Corporation of America, 575 Technology Square, Cambridge, Massachusetts 02139.

[LINCOLN]

ESD-TR-78-64 Semiannual Technical Summary, Seismic Discrimination, 31 March 1978, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02173.

[MARILL & STERN]

Marill, Thomas; and Stern, D.H. "The Datacomputer: A Network Data Utility", Proceedings AFIPS National Computer Conference, AFIPS Press, Vol. 44, 1975.

[SATTLEY a]

Joanne Z. Sattley, "Report on the Implementation and Test of SWF-D: The Signal Waveform File Demon", Technical Report CCA-79-09, 31 January 1979, Computer Corporation of America, 575 Technology Square, Cambridge, Massachusetts 02139.

[SATTLEY b]

Joanne Z.Sattley, "Program Listings for SWF-D: The Signal Waveform File Demon", Technical Report CCA-79-10, 31 January 1979, Computer Corporation of America, 575 Technology Square, Cambridge, Massachusetts 02139.

[SATTLEY c]

Joanne Z. Sattley, "MARS - A Message Archiving and Retrieval Service", 8 January 1978, ARPA Network Working Group Request for Comments #744, Network Information Center #42857.

[TELEDYNE]

VELANET ESF/SWF PROCESSING, Joseph Greenhalgh, Design Draft, February 1, 1978, Teledyne Geotech, Alexandria Laboratories, P. O. Box 334, Alexandria, Virginia 22313.